



Taking a Stratified Sample from a Data Set

Task:

Stratified Samples Solution

A stratified random sample is used when distinct groups or strata are identified as important within a data set. By using this sampling method, each group is represented in the sample in the same proportion as in the population from which the sample is drawn.

For example if I decide to collect a sample of 20 from a population of 600 boys and 400 girls then each group in the sample will be represented in the same fraction by:

Boys $600/1000 = 60\%$ and girls $400/1000 = 40\%$. Therefore, in the sample of 20 you would need to randomly select 60% of $20 = 12$ boys, and 40% of $20 = 8$ girls.

Look at the following data showing the number of students who filled in a questionnaire in the 2008 CensusAtSchool project.

Year Level	ACT	NSW	NT	QLD	SA	TAS	VIC	WA	Year Totals
Year 4 or below	12	919	5	77	176	152	549	176	2066
Year 5	171	2321	19	208	350	419	2070	587	6145
Year 6	176	3235	29	452	782	654	2674	590	8592
Year 7	57	1379	25	801	1007	376	1411	1272	6328
Year 8	76	2131	17	860	1329	536	1095	954	6998
Year 9	33	1857	18	1016	1244	518	950	912	6548
Year 10	30	1100	17	677	1054	354	502	914	4648
Year 11	49	447	41	343	536	194	205	159	1974
Year 12	27	105	3	227	253	38	49	78	780
Other	7	20	4	10	14	8	29	18	110
Total	638	13514	178	4671	6745	3249	9534	5660	44,189

Table 1: Number of students who completed the CensusAtSchool questionnaire by year level and State

1. Name two ways that this data has been organised that could be used to make stratified random samples. *State and Year Level*
2. A total of 44,189 students participated in the 2008 CensusAtSchool questionnaire. For a stratified sample based on the State for each student, calculate the numbers in the sample for each State if the sample size is to be 100.



- a) Give you answers accurate to 2 decimal places.
 b) Round this number to the nearest whole number.
 c) Check that you total for the whole number calculations add up to 100.

Year Level	ACT	NSW	NT	QLD	SA	TAS	VIC	WA	Year Totals
total	638	13514	178	4671	6745	3249	9534	5660	44189
2 dp	<i>1.44</i>	<i>30.58</i>	<i>0.40</i>	<i>10.57</i>	<i>15.26</i>	<i>7.35</i>	<i>21.58</i>	<i>12.81</i>	100
nearest whole number	<i>1</i>	<i>31</i>	<i>0</i>	<i>11</i>	<i>15</i>	<i>7</i>	<i>22</i>	<i>13</i>	100

Table 2: Number of students who completed the CensusAtSchool questionnaire by State

- 3.** For a sample size of 50 calculate the number needed if you formed stratified sample according to the Year Level of each student.
- d) Give you answers accurate to 2 decimal places.
 e) Round this number to the nearest whole number.
 f) Check that your total for the whole number calculations add up to 50.

Year Level	year totals	2 dp	nearest whole number
Year 4 or below	2066	<i>2.34</i>	<i>2</i>
Year 5	6145	<i>6.95</i>	<i>7</i>
Year 6	8592	<i>9.72</i>	<i>10</i>
Year 7	6328	<i>7.16</i>	<i>7</i>
Year 8	6998	<i>7.92</i>	<i>8</i>
Year 9	6548	<i>7.41</i>	<i>7</i>
Year 10	4648	<i>5.26</i>	<i>5</i>
Year 11	1974	<i>2.23</i>	<i>2</i>
Year 12	780	<i>0.88</i>	<i>1</i>
Other	110	<i>0.12</i>	<i>0</i>
Total	44189	50	50

Table 3: Number of students who completed the CensusAtSchool questionnaire by Year Level

Extension:

CensusAtSchool Random Sampler <http://www.cas.abs.gov.au/cgi-local/cassampler.pl>
 Access other similar lessons at <http://www.abs.gov.au/censusatschool>



For each of the following, based on a co-educational school population of 1000 students, decide whether or not you think that it would be warranted to obtain a stratified random sample rather than a simple random sample. If so, what strata would you include and why?

- 1.** The junior school coordinator wants to know if junior students want the lunchtime equipment program to continue.

Yes, views might differ according to year level.

- 2.** The canteen wants to sample students to see whether students like the food on their current menu.

Yes, views might differ according to year level and gender

- 3.** The Principal wants students' opinion on whether to include a blazer in the uniform.

No

- 4.** The Sport teachers want to know whether to include aerobics, soccer or surfing on their elective program.

Yes, views might differ according to gender

- 5.** The library isn't sure whether to continue its subscription to "Girlfriend" magazine.

Yes, views might differ according to gender